

“Express Mail” Mailing Label No. **EL917901516US**

**PATENT APPLICATION
ATTORNEY DOCKET NO. SUN-P6430-SPL**

5

10 **METHOD AND APPARATUS FOR SOLVING
AN EQUALITY CONSTRAINED GLOBAL
OPTIMIZATION PROBLEM**

15 **Inventors:** G. William Walster and Eldon R. Hansen

Related Application

20 The subject matter of this application is related to the subject matter in a co-pending non-provisional application by the same inventors as the instant application and filed on the same day as the instant application entitled, "Applying Term Consistency to an Equality Constrained Interval Global Optimization Problem," having serial number TO BE ASSIGNED, and filing date 13 December 2001 (Attorney Docket No. SUN-P6445-SPL).

25

BACKGROUND

Field of the Invention

30 The present invention relates to performing arithmetic operations on interval operands within a computer system. More specifically, the present

invention relates to a method and an apparatus for using a computer system to solve a global optimization problem including equality constraints with interval arithmetic.

5 Related Art

Rapid advances in computing technology make it possible to perform trillions of computational operations each second. This tremendous computational speed makes it practical to perform computationally intensive tasks as diverse as predicting the weather and optimizing the design of an aircraft
10 engine. Such computational tasks are typically performed using machine-representable floating-point numbers to approximate values of real numbers. (For example, see the Institute of Electrical and Electronics Engineers (IEEE) standard 754 for binary floating-point numbers.)

In spite of their limitations, floating-point numbers are generally used to
15 perform most computational tasks.

One limitation is that machine-representable floating-point numbers have a fixed-size word length, which limits their accuracy. Note that a floating-point number is typically encoded using a 32, 64 or 128-bit binary number, which means that there are only 2^{32} , 2^{64} or 2^{128} possible symbols that can be used to
20 specify a floating-point number. Hence, most real number values can only be approximated with a corresponding floating-point number. This creates estimation errors that can be magnified through even a few computations, thereby adversely affecting the accuracy of a computation.

A related limitation is that floating-point numbers contain no information
25 about their accuracy. Most measured data values include some amount of error that arises from the measurement process itself. This error can often be quantified as an accuracy parameter, which can subsequently be used to determine the

accuracy of a computation. However, floating-point numbers are not designed to keep track of accuracy information, whether from input data measurement errors or machine rounding errors. Hence, it is not possible to determine the accuracy of a computation by merely examining the floating-point number that results from

5 the computation.

Interval arithmetic has been developed to solve the above-described problems. Interval arithmetic represents numbers as intervals specified by a first (left) endpoint and a second (right) endpoint. For example, the interval $[a, b]$, where $a < b$, is a closed, bounded subset of the real numbers, R , which includes a 10 and b as well as all real numbers between a and b . Arithmetic operations on interval operands (interval arithmetic) are defined so that interval results always contain the entire set of possible values. The result is a mathematical system for rigorously bounding numerical errors from all sources, including measurement data errors, machine rounding errors and their interactions. (Note that the first 15 endpoint normally contains the “infimum”, which is the largest number that is less than or equal to each of a given set of real numbers. Similarly, the second endpoint normally contains the “supremum”, which is the smallest number that is greater than or equal to each of the given set of real numbers.)

One commonly performed computational operation is to perform equality 20 constrained global optimization to find a global minimum of a nonlinear objective function subject to nonlinear equality constraints of the form $q_i(\mathbf{x}) = 0$ ($i=1, \dots, r$). This can be accomplished by deleting boxes, or parts of boxes that do not satisfy one or more equality constraints, or by unconditionally deleting boxes that cannot contain the global minimum f^* .

25 To delete part or all of a subbox that fails to satisfy one or more equality constraints, term consistency and the interval Newton method can be used in combination.

To unconditionally delete part or all of a box that cannot contain the equality constrained global minimum, the smallest upper bound f_{bar} so far computed on the global minimum f^* can be used.

In the presence of equality constraints, it can be difficult to identify a feasible point x that is guaranteed to satisfy all the given equality constraints.

5 However, it is relatively easy to prove that a box X contains a feasible point. Given such a box, $sup(f(X))$ provides a value of f_{bar} . As this process progresses, f_{bar} is updated using $f_{bar} = \min(f_{bar}, sup(f(X)))$, given that the box X has been proved to contain a feasible point.

10 Given an f_{bar} value, the f_{bar} criterion is applied as an inequality constraint to delete parts or all of a given subbox.

If f_{bar} is the smallest upper bound so far computed on f , then any point x for which $f(x) > f_{bar}$ can be deleted. Similarly, any box X can be deleted if $inf(f(X)) > f_{bar}$.

15 In addition to solving for boxes with feasible points and the f_{bar} criterion, the John conditions can be solved to delete parts or all of subboxes and to prove that a given subbox contains a feasible point. (The John conditions are described in “Global Optimization Using Interval Analysis” by Eldon R. Hansen, Marcel Dekker, Inc., 1992.)

20 Solving the John conditions works best “in the small” when the objective function f is approximately quadratic and satisfied constraints are approximately linear. For large boxes containing multiple stationary points, solving the John conditions might not succeed in deleting much of a given box. In this case the box is split into two or more sub-boxes that are then processed independently. By

25 this mechanism *all* the equality constrained global minima of a nonlinear objective function can be found.

One problem is applying this procedure to large n -dimensional interval vectors (or boxes) that contain multiple local minima. In this case, the process of splitting in n -dimensions can lead to exponential growth in the number of boxes to process.

5 It is well known that this problem (and even the problem of computing “sharp” bounds on the range of a function of n -variables over an n -dimensional box) is an “NP-hard” problem. In general, NP-hard problems require an exponentially increasing amount of work to solve as n , the number of independent variables, increases.

10 Because NP-hardness is a worst-case property and because many practical engineering and scientific problems have relatively simple structure, one problem is to use this simple structure of real problems to improve the efficiency of interval equality constrained global optimization algorithms.

Hence, what is needed is a method and an apparatus for using the structure
15 of a nonlinear objective function to improve the efficiency of interval equality constrained global optimization software. To this end, what is needed is a method and apparatus that efficiently deletes “large” boxes or parts of large boxes that using the interval Newton method to solve the John conditions can only split.

20

SUMMARY

The present invention combines various methods to speed up the process of bounding all the equality constrained global minima of a nonlinear objective function. The combined method uses the structure of the objective function and equality constraints to efficiently delete parts or all of large boxes that would
25 otherwise have to be split. One embodiment of the present invention provides a system that solves a global optimization problem specified by a function f and a set of equality constraints $q_i(\mathbf{x}) = 0$ ($i=1, \dots, r$), wherein f is a scalar function of a

vector $\mathbf{x} = (x_1, x_2, x_3, \dots, x_n)$. During operation, the system receives a representation of the function f and the set of equality constraints and stores the representation in a memory. Next, the system performs an interval equality constrained global optimization process to compute guaranteed bounds on a

5 globally minimum value of the function $f(\mathbf{x})$ subject to the set of equality constraints. During this process, the system applies term consistency to a set of relations associated with the interval equality constrained global optimization problem over a subbox \mathbf{X} , and excludes any portion of the subbox \mathbf{X} that violates any of these relations. It also applies box consistency to the set of relations, and

10 excludes any portion of the subbox \mathbf{X} that violates any of the relations. Finally, the system performs an interval Newton step for the interval equality constrained global optimization problem over the subbox \mathbf{X} .

In a variation on the above embodiment, the system applies term consistency to the set of equality constraints $q_i(\mathbf{x}) = 0$ ($i=1,\dots,r$) over the subbox

15 \mathbf{X} .

In a variation on the above embodiment, the system applies box consistency to the set of equality constraints $q_i(\mathbf{x}) = 0$ ($i=1,\dots,r$) over the subbox \mathbf{X} .

In a variation on the above embodiment, while performing the interval

20 equality constrained global optimization process, the system keeps track of a least upper bound f_{bar} of the function $f(\mathbf{x})$, and removes from consideration any subbox for which $\inf(f(\mathbf{X})) > f_{\text{bar}}$. In this variation, the system applies term consistency to the f_{bar} inequality $f(\mathbf{x}) \leq f_{\text{bar}}$ over the subbox \mathbf{X} .

In a further variation, the system applies box consistency to the f_{bar}

25 inequality $f(\mathbf{x}) \leq f_{\text{bar}}$ over the subbox \mathbf{X} .

In a variation on the above embodiment, while performing the interval

equality constrained global optimization process, the system preconditions the set

of equality constraints through multiplication by an approximate inverse matrix \mathbf{B} to produce a set of preconditioned equality constraints. In this variation, the system applies term consistency to the set of preconditioned equality constraints over the subbox \mathbf{X} .

5 In a further variation, the system applies box consistency to the set of preconditioned equality constraints over the subbox \mathbf{X} .

In a variation on the above embodiment, the system performs the interval Newton step on the John conditions. In a further variation, prior to performing the interval Newton step on the John conditions, the system performs a linearization
10 test to determine whether to perform the Newton step on the John conditions.

In a variation on the above embodiment, while performing the interval global optimization process, the system evaluates a first termination condition. This first termination condition is TRUE if the width of the subbox \mathbf{X} is less than a pre-specified value, ε_X , and the width of $f(\mathbf{X})$ is less than a pre-specified value,
15 ε_F . If the first termination condition is TRUE, the system terminates further splitting of the subbox \mathbf{X} .

In a variation on the above embodiment, while applying term consistency the system symbolically manipulates an equation to solve for a term, $g(x_j)$, thereby producing a modified equation $g(x_j) = h(\mathbf{x})$, wherein the term $g(x_j)$ can be
20 analytically inverted to produce an inverse function $g^{-1}(\mathbf{y})$. Next, the system substitutes the subbox \mathbf{X} into the modified equation to produce the equation $g(X'_j) = h(\mathbf{X})$, and solving for $X'_j = g^{-1}(h(\mathbf{X}))$. Finally, the system intersects X'_j with the
interval X_j to produce a new subbox \mathbf{X}^+ , wherein the new subbox \mathbf{X}^+ contains all
solutions of the equation within the subbox \mathbf{X} , and wherein the size of the new
25 subbox \mathbf{X}^+ is less than or equal to the size of the subbox \mathbf{X} .

Note that for a given function of n -variables $f(\mathbf{x})$ there are different ways to analytically solve for a component x_j of the vector \mathbf{x} . For example, one can write

$f(\mathbf{x}) = g(x_j) - h(\mathbf{x})$, where $g(x_j)$ is a term in f for which it is possible to solve $g(x_j) = y$ for any x_j using $g^{-1}(y)$. For each of these rearrangements, if a given interval box \mathbf{X} is used as an argument of h , then the new interval X_j^+ for the j -th component of \mathbf{X} , is guaranteed to be at least as narrow as the original, X_j .

5

$$X_j^+ = X_j \cap X'_j \text{ where } X'_j = g^{-1}(h(\mathbf{X})).$$

This process is then iterated using different terms g of the function f . After reducing any element X_j of the box \mathbf{X} to X_j^+ , the reduced value can be used in \mathbf{X} thereafter to speed up the reduction process using other component functions and terms thereof if f is a component of the vector function \mathbf{f} .

Hereafter, the notation $g(x_j)$ for a term of the function $f(\mathbf{x})$ implicitly represents any term of any component function. This eliminates the need for additional subscripts that do not add clarity to the exposition.

15

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1 illustrates a computer system in accordance with an embodiment of the present invention.

FIG. 2 illustrates the process of compiling and using code for interval computations in accordance with an embodiment of the present invention.

FIG. 3 illustrates an arithmetic unit for interval computations in accordance with an embodiment of the present invention.

FIG. 4 is a flow chart illustrating the process of performing an interval computation in accordance with an embodiment of the present invention.

25

FIG. 5 illustrates four different interval operations in accordance with an embodiment of the present invention.

FIG. 6 is a flow chart illustrating the process of finding an interval solution to a nonlinear equation in accordance with an embodiment of the present invention.

5 FIG. 7A presents a first portion of a flow chart illustrating the process of solving an interval global optimization problem with equality constraints in accordance with an embodiment of the present invention.

FIG. 7B presents a second portion of a flow chart illustrating the process of solving an interval global optimization problem with equality constraints in accordance with an embodiment of the present invention.

10 FIG. 7C presents a third portion of a flow chart illustrating the process of solving an interval global optimization problem with equality constraints in accordance with an embodiment of the present invention.

15 FIG. 7D presents a fourth portion of a flow chart illustrating the process of solving an interval global optimization problem with equality constraints in accordance with an embodiment of the present invention.

FIG. 7E presents a fifth portion of a flow chart illustrating the process of solving an interval global optimization problem with equality constraints in accordance with an embodiment of the present invention.

20 **DETAILED DESCRIPTION**

The following description is presented to enable any person skilled in the art to make and use the invention, and is provided in the context of a particular application and its requirements. Various modifications to the disclosed embodiments will be readily apparent to those skilled in the art, and the general principles defined herein may be applied to other embodiments and applications without departing from the spirit and scope of the present invention. Thus, the
25

present invention is not limited to the embodiments shown, but is to be accorded the widest scope consistent with the principles and features disclosed herein.

The data structures and code described in this detailed description are typically stored on a computer readable storage medium, which may be any device or medium that can store code and/or data for use by a computer system. This includes, but is not limited to, magnetic and optical storage devices such as disk drives, magnetic tape, CDs (compact discs) and DVDs (digital versatile discs or digital video discs), and computer instruction signals embodied in a transmission medium (with or without a carrier wave upon which the signals are modulated).
5 For example, the transmission medium may include a communications network, such as the Internet.
10

Computer System

FIG. 1 illustrates a computer system 100 in accordance with an embodiment of the present invention. As illustrated in FIG. 1, computer system 100 includes processor 102, which is coupled to a memory 112 and a peripheral bus 110 through bridge 106. Bridge 106 can generally include any type of circuitry for coupling components of computer system 100 together.
15

Processor 102 can include any type of processor, including, but not limited to, a microprocessor, a mainframe computer, a digital signal processor, a personal organizer, a device controller and a computational engine within an appliance. Processor 102 includes an arithmetic unit 104, which is capable of performing computational operations using floating-point numbers.
20

Processor 102 communicates with storage device 108 through bridge 106 and peripheral bus 110. Storage device 108 can include any type of non-volatile storage device that can be coupled to a computer system. This includes, but is not
25

limited to, magnetic, optical, and magneto-optical storage devices, as well as storage devices based on flash memory and/or battery-backed up memory.

Processor 102 communicates with memory 112 through bridge 106. Memory 112 can include any type of memory that can store code and data for execution by processor 102. As illustrated in FIG. 1, memory 112 contains computational code for intervals 114. Computational code 114 contains instructions for the interval operations to be performed on individual operands, or interval values 115, which are also stored within memory 112. This computational code 114 and these interval values 115 are described in more detail below with reference to FIGS. 2-5.

Note that although the present invention is described in the context of computer system 100 illustrated in FIG. 1, the present invention can generally operate on any type of computing device that can perform computations involving floating-point numbers. Hence, the present invention is not limited to the computer system 100 illustrated in FIG. 1.

Compiling and Using Interval Code

FIG. 2 illustrates the process of compiling and using code for interval computations in accordance with an embodiment of the present invention. The system starts with source code 202, which specifies a number of computational operations involving intervals. Source code 202 passes through compiler 204, which converts source code 202 into executable code form 206 for interval computations. Processor 102 retrieves executable code 206 and uses it to control the operation of arithmetic unit 104.

Processor 102 also retrieves interval values 115 from memory 112 and passes these interval values 115 through arithmetic unit 104 to produce results 212. Results 212 can also include interval values.

Note that the term “compilation” as used in this specification is to be construed broadly to include pre-compilation and just-in-time compilation, as well as use of an interpreter that interprets instructions at run-time. Hence, the term “compiler” as used in the specification and the claims refers to pre-compilers, 5 just-in-time compilers and interpreters.

Arithmetic Unit for Intervals

FIG. 3 illustrates arithmetic unit 104 for interval computations in more detail accordance with an embodiment of the present invention. Details regarding 10 the construction of such an arithmetic unit are well known in the art. For example, see U.S. Patent Nos. 5,687,106 and 6,044,454. Arithmetic unit 104 receives intervals 302 and 312 as inputs and produces interval 322 as an output.

In the embodiment illustrated in FIG. 3, interval 302 includes a first floating-point number 304 representing a first endpoint of interval 302, and a 15 second floating-point number 306 representing a second endpoint of interval 302. Similarly, interval 312 includes a first floating-point number 314 representing a first endpoint of interval 312, and a second floating-point number 316 representing a second endpoint of interval 312. Also, the resulting interval 322 includes a first floating-point number 324 representing a first endpoint of interval 20 322, and a second floating-point number 326 representing a second endpoint of interval 322.

Note that arithmetic unit 104 includes circuitry for performing the interval operations that are outlined in FIG. 5. This circuitry enables the interval operations to be performed efficiently.

25 However, note that the present invention can also be applied to computing devices that do not include special-purpose hardware for performing interval operations. In such computing devices, compiler 204 converts interval operations

into a executable code that can be executed using standard computational hardware that is not specially designed for interval operations.

FIG. 4 is a flow chart illustrating the process of performing an interval computation in accordance with an embodiment of the present invention. The system starts by receiving a representation of an interval, such as first floating-point number 304 and second floating-point number 306 (step 402). Next, the system performs an arithmetic operation using the representation of the interval to produce a result (step 404). The possibilities for this arithmetic operation are described in more detail below with reference to FIG. 5.

10

Interval Operations

FIG. 5 illustrates four different interval operations in accordance with an embodiment of the present invention. These interval operations operate on the intervals X and Y . The interval X includes two endpoints,

- \underline{x} denotes the lower bound of X , and
- \bar{x} denotes the upper bound of X .

15

The interval X is a closed subset of the extended (including $-\infty$ and $+\infty$) real numbers R^* (see line 1 of FIG. 5). Similarly the interval Y also has two endpoints and is a closed subset of the extended real numbers R^* (see line 2 of FIG. 5).

Note that an interval is a point or degenerate interval if $X = [\underline{x}, \bar{x}]$. Also note that the left endpoint of an interior interval is always less than or equal to the right endpoint. The set of extended real numbers, R^* is the set of real numbers, R , extended with the two ideal points negative infinity and positive infinity:

$$R^* = R \cup \{-\infty\} \cup \{+\infty\}.$$

25

In the equations that appear in FIG. 5, the up arrows and down arrows indicate the direction of rounding in the next and subsequent operations. Directed rounding (up or down) is applied if the result of a floating-point operation is not machine-representable.

5 The addition operation $X + Y$ adds the left endpoint of X to the left endpoint of Y and rounds down to the nearest floating-point number to produce a resulting left endpoint, and adds the right endpoint of X to the right endpoint of Y and rounds up to the nearest floating-point number to produce a resulting right endpoint.

10 Similarly, the subtraction operation $X - Y$ subtracts the right endpoint of Y from the left endpoint of X and rounds down to produce a resulting left endpoint, and subtracts the left endpoint of Y from the right endpoint of X and rounds up to produce a resulting right endpoint.

15 The multiplication operation selects the minimum value of four different terms (rounded down) to produce the resulting left endpoint. These terms are: the left endpoint of X multiplied by the left endpoint of Y ; the left endpoint of X multiplied by the right endpoint of Y ; the right endpoint of X multiplied by the left endpoint of Y ; and the right endpoint of X multiplied by the right endpoint of Y . This multiplication operation additionally selects the maximum of the same four terms (rounded up) to produce the resulting right endpoint.

20 Similarly, the division operation selects the minimum of four different terms (rounded down) to produce the resulting left endpoint. These terms are: the left endpoint of X divided by the left endpoint of Y ; the left endpoint of X divided by the right endpoint of Y ; the right endpoint of X divided by the left endpoint of Y ; and the right endpoint of X divided by the right endpoint of Y . This division operation additionally selects the maximum of the same four terms (rounded up) to produce the resulting right endpoint. For the special case where the interval Y

includes zero, X/Y is an exterior interval that is nevertheless contained in the interval R^* .

Note that the result of any of these interval operations is the empty interval if either of the intervals, X or Y , are the empty interval. Also note, that in one embodiment of the present invention, extended interval operations never cause undefined outcomes, which are referred to as “exceptions” in the IEEE 754 standard.

Term Consistency

FIG. 6 is a flow chart illustrating the process of solving a nonlinear equation through interval arithmetic and term consistency in accordance with an embodiment of the present invention. The system starts by receiving a representation of a nonlinear equation $f(\mathbf{x}) = 0$ (step 602), as well as a representation of an initial box with \mathbf{X} with X_j an element of \mathbf{X} (step 604). Next, the system symbolically manipulates the equation $f(\mathbf{x}) = 0$ into a form $g(x_j) - h(\mathbf{x}) = 0$, wherein the term $g(x_j)$ can be analytically inverted to produce an inverse function $g^{-1}(y)$ (step 606).

Next, the system substitutes the initial box \mathbf{X} into $h(\mathbf{X})$ to produce the equation $g(X'_j) = h(\mathbf{X})$ (step 608). The system then solves for $X'_j = g^{-1}(h(\mathbf{X}))$ (step 610). The resulting interval X'_j is then intersected with the initial interval X_j to produce a new interval X_j^+ (step 612).

At this point, if X_j^+ is empty, the system can terminate. Otherwise, the system can perform further processing. This further processing involves saving X_j by setting $X^{(0)}_j = X_j$ and also, setting $X_j = X_j^+$ (step 614). Next, if $w(X^{(0)})$ is sufficiently reduced at step 616, the system returns to step 606 for another iteration of term consistency on another term g of $f(\mathbf{x})$. Otherwise, the system terminates.

Examples of Applying Term Consistency

For example, suppose $f(x) = x^2 - x + 6$. We can define $g(x) = x^2$ and $h(x) = x - 6$. Let $X = [-10, 10]$. The procedural step is $(X')^2 = X - 6 = [-16, 4]$. Since $(X')^2$ must be non-negative, we replace this interval by $[0, 4]$. Solving for X' , we obtain $X' = \pm [0, 2]$. In replacing the range of $h(x)$ (i.e., $[-16, 4]$) by non-negative values, we have excluded that part of the range $h(x)$ that is not in the domain of $g(x) = x^2$.

Suppose that we reverse the roles of g and h and use the iterative step
 10 $h(X') = g(X)$. That is $X' - 6 = X^2$. We obtain $X' = [6,106]$. Intersecting this result with the interval $[-10,10]$, of interest, we obtain $[6,10]$. This interval excludes the set of values for which the range of $g(X)$ is not in the intersection of the domain of $h(X)$ with X .

Combining these results, we conclude that any solution of
 15 $f(X) = g(X) - h(X) = 0$ that occurs in $X = [-10, 10]$ must be in both $[-2, 2]$ and $[6, 10]$. Since these intervals are disjoint, there can be no solution in $[-10, 10]$.

In practice, if we already reduced the interval from $[-10,10]$ to $[-2,2]$ by solving for g , we use the narrower interval as input when solving for h .

This example illustrates the fact that it can be advantageous to solve a given equation for more than one of its terms. The order in which terms are chosen affects the efficiency. Unfortunately, it is not known how best to choose the best order.

Also note that there can be many choices for $g(x)$. For example, suppose we use term consistency to narrow the interval bound X on a solution of
 25 $f(x) = ax^4 + bx + c = 0$. We can let $g(x) = bx$ and compute $X' = -(ax^4 + c)/b$ or we can let $g(x) = ax^4$ and compute $X' = \pm [-(bx+c)/a]^{1/4}$. We can also separate x^4 into $x^2 * x^2$ and solve for one of the factors $X' = \pm [-(bx+c)/(ax^2)]^{1/2}$.

In the multidimensional case, we may solve for a term involving more than one variable. We then have a two-stage process. For example, suppose we solve for the term $1/(x+y)$ from the function $f(x,y) = 1/(x+y) - h(x,y) = 0$. Let $x \in X = [1,2]$ and $y \in Y = [0.5,2]$. Suppose we find that $h(X,Y) = [0.5,1]$. Then 5 $1/(x+y) \in [0.5,1]$ so $x+y \in [1,2]$. Now we replace y by $Y = [0.5,2]$ and obtain the bound $[-1,1.5]$ on X . Intersecting this interval with the given bound $X = [1,2]$ on x , we obtain the new bound $X' = [1,1.5]$.

We can use X' to get a new bound on h ; but this may require extensive computing if h is a complicated function; so suppose we do not. Suppose that we 10 do, however, use this bound on our intermediate result $x + y = [1,2]$. Solving for y as $[1,2] - X'$, we obtain the bound $[-0.5,1]$. Intersecting this interval with Y , we obtain the new bound $Y' = [0.5,1]$ on y . Thus, we improve the bounds on both x and y by solving for a single term of f .

The point of these examples is to show that term consistency can be used 15 in many ways both alone and in combination with the interval Newton algorithm to improve the efficiency with which roots of a single nonlinear equation can be computed. The same is true for systems of nonlinear equations.

Equality Constrained Interval Global Optimization

FIGs. 7A-7E collectively present a flow chart illustrating the process of 20 solving a global optimization problem with equality constraints in accordance with an embodiment of the present invention. Generally, we seek a solution in a single box specified by the user. However, any number of boxes can be specified.

The boxes can be disjoint or overlap. However, if they overlap, a 25 minimum at a point that is common to more than one box is separately found as a solution in each box containing it. In this case, computing effort is wasted. If the user does not specify an initial box or boxes, we use a default box. The process

described below finds the global minimum in the set of points formed by the set of boxes. We assume these initial boxes are placed in a list L_I of boxes to be processed.

Suppose the user of the process knows a point x_{bar} that is guaranteed to
5 be feasible. If so, we use this point to compute an initial upper bound f_{bar} on the global minimum f^* . If x_{bar} cannot be represented exactly on the computer, the system forms a representable interval vector \mathbf{X} containing x_{bar} . We evaluate $f(\mathbf{X})$ and obtain [lower bound $f(\mathbf{X})$, upper bound $f(\mathbf{X})$]. Even if rounding and/or dependence are such that \mathbf{X} cannot be numerically proven to be certainly feasible,
10 we rely upon the user and assume that \mathbf{X} contains a feasible point. Therefore, we set f_{bar} equal to the upper bound of $f(\mathbf{X})$.

Also the user may know an upper bound f_{bar} on f^* even though he may not know where (or even if) f takes on such a value in the feasible region defined by the equality constraints. If so, we set f_{bar} equal to this known bound. If the
15 known bound is not representable on the computer, the system rounds the value up to a larger value that is representable.

If no feasible point is known and no upper bound on f^* is known, we set $f_{\text{bar}} = +\infty$. The user must specify a box size tolerance ε_X and a function width tolerance ε_F .

20 In the system, nonlinear functions are often linearized over a box \mathbf{X} using Taylor expansion. However, use of linearization is generally ineffective if \mathbf{X} is wide. Four different sub-procedures in the system use linearization. The system uses a “linearization test” to decide if a given sub-procedure should be used for a given box. Each of the four sub-procedures uses a separate test of the same kind.
25 In each case, a criterion for “success” is defined. The symbol w_R denotes the width of the largest box for which success was achieved. The symbol w_I denotes the width of the smallest box for which success was not achieved. A given sub-

procedure is applied for a box \mathbf{X} whenever $w(\mathbf{X}) \leq (w_R + w_I)/2$. For each sub-procedure, the system initially sets $w_R = 0$ and $w_I = w(\mathbf{X}^{(0)})$, where $\mathbf{X}^{(0)}$ is the initial box. In addition, the system specifies a bound f_bar if one is known. Note that the four sub-procedures referred to above are: (1) Newton applied to the gradient; (2) Newton applied to the John conditions; (3) linearization of the constraints; and (4) linearization of $f(\mathbf{x}) \leq f_bar$.

5 The steps of the process are performed in the order given except as indicated by branching.

For each initial box \mathbf{X} in the list L_I , the system evaluates $f(\mathbf{X})$ (step 701).

10 Next, if $f_bar < \infty$, the system deletes any box \mathbf{X} from L_I for which the lower bound of $f(\mathbf{X}) > f_bar$. This can be done while applying term consistency (step 702).

15 If L_I is empty, the system goes to step 745. Otherwise, the system finds the box \mathbf{X} in L_I for which the lower bound of $f(\mathbf{X})$ is smallest. For later reference, the system calls this box $\mathbf{X}^{(1)}$. This box is processed next. The system deletes $\mathbf{X}^{(1)}$ from L_I (step 703).

Next, if $w[f(\mathbf{X})] > \varepsilon_F$, the system goes to step 706 (step 704).

Then, if $w(\mathbf{X}) \leq \varepsilon_X$, the system puts \mathbf{X} in list L_2 and goes to step 703 (step 705).

20 Next, the system applies term consistency to the constraint equations $q_i(\mathbf{x}) = 0$ ($i=1,\dots,r$) for $i=1,\dots,r$. If it is proved that there is no point in \mathbf{X} that satisfies any one of the constraints, the system goes to step 703 (step 706).

If $sup(f(\mathbf{X})) \leq f_bar$, the system goes to step 711 (step 707).

25 Next, if $w[f(\mathbf{X})] \leq \varepsilon_F$ and $w(\mathbf{X}) \leq \varepsilon_X$, the system puts \mathbf{X} in list L_2 and goes to step 703 (step 708).

Then, if $\mathbf{X}^{(1)}$ (as defined in step 703) has been sufficiently reduced, the system puts \mathbf{X} in the list L_I and goes to Step 703 (step 709). We say that a box \mathbf{X}

is sufficiently reduced if any component of \mathbf{X} is reduced by an amount that is at least a fraction (say 0.25) of the width of the widest component of \mathbf{X} .

If $f[m(\mathbf{X})] \geq f_bar$, the system goes to step 713 (step 710).

For later reference, the system calls the current box $\mathbf{X}^{(2)}$. The system then

5 uses a linearization test to decide whether to linearize and “solve” the inequality $f(\mathbf{x}) \leq f_bar$. Let w_I denote the width of the smallest box for which \mathbf{M}^I is irregular. Let w_R denote the width of the largest box for which \mathbf{M}^I has been found to be regular. Initially, set $w_I = w(\mathbf{X}^{(0)})$ and $w_R = 0$. We use w_I and w_R to decide whether to generate and attempt to solve $\mathbf{g}(\mathbf{x}) + \mathbf{J}(\mathbf{x}, \mathbf{X})(\mathbf{y} - \mathbf{x}) = 0$ for a given box \mathbf{X} .

10 We do so if $w(\mathbf{X}) \leq (w_I + w_R)/2$. (see Chapter 12 of Global Optimization Using Interval Analysis by Eldon R. Hansen, Marcel Dekker, Inc., 1992) (step 711).

If $\text{sup}(f(\mathbf{X})) \leq f_bar$, the system goes to step 714 (step 712).

Next, the system applies term consistency to the relation $f(\mathbf{x}) \leq f_bar$. If the result is empty, the system goes to step 703 (step 713).

15 If $w[f(\mathbf{X})] \leq \varepsilon_F$ and $w(\mathbf{X}) \leq \varepsilon_X$, the system puts \mathbf{X} in list L_2 and goes to step 703 (step 714).

If $\mathbf{X}^{(1)}$ (as defined in 703) has been sufficiently reduced, the system puts \mathbf{X} in L_1 and goes to step 703 (step 715).

Next, the system applies box consistency to the constraint equations

20 $q_i(\mathbf{x}) = 0$ ($i=1, \dots, r$) for $i=1, \dots, r$. If it is proved that there is no point in \mathbf{X} that satisfies any one of the constraints, the system goes to step 703 (step 716).

If $\text{inf}(f[m(\mathbf{X})]) \geq f_bar$, the system goes to step 722 (step 717).

If the current box is the same box $\mathbf{X}^{(2)}$ defined in step 711, the system goes to step 720 (step 718).

25 Next, the system again uses the procedure in step 711 to try to reduce the upper bound f_bar (step 719).

If $\text{sup}(f(\mathbf{X})) \leq f_bar$, the system goes to step 722 (step 720).

Next, the system applies box consistency to the relation $f(\mathbf{x}) \leq f_bar$. If the result is empty, the system goes to step 703 (step 721).

If $w[f(\mathbf{X})] \leq \varepsilon_F$ and $w(\mathbf{X}) \leq \varepsilon_X$, the system puts \mathbf{X} in list L_2 and goes to step 703 (step 722).

5 Next, if $\mathbf{X}^{(1)}$ (as defined in step 703) has been sufficiently reduced, the system puts \mathbf{X} in the list L_1 and goes to step 703 (step 723).

If $inf(f[m(\mathbf{X})]) \geq f_bar$, the system goes to step 713 (step 724).

Next, the system denotes the current box by $\mathbf{X}^{(3)}$, and applies a linear method involving a Taylor expansion to try to reduce $\mathbf{X}^{(3)}$ using $f(\mathbf{x}) \leq f_bar$. If the result is empty, the system goes to step 703 (step 725).

10 Then, if $w[f(\mathbf{X})] \leq \varepsilon_F$ and $w(\mathbf{X}) \leq \varepsilon_X$, the system puts \mathbf{X} in list L_2 and goes to step 703 (step 726).

If $\mathbf{X}^{(3)}$ as defined in step 726 was not sufficiently reduced in the single step 725, the system goes to step 730.

15 Next, the system applies a quadratic method in which f is expanded through quadratic terms to try to reduce the current box using $f(\mathbf{x}) \leq f_bar$. If the result is empty, the system goes to step 703 (step 728).

If $w[f(\mathbf{X})] \leq \varepsilon_F$ and $w(\mathbf{X}) \leq \varepsilon_X$, the system puts \mathbf{X} in list L_2 and goes to step 703 (step 729).

20 If $\mathbf{X}^{(1)}$ (as defined in step 703) has been sufficiently reduced, the system puts \mathbf{X} in L_1 and goes to step 703 (step 730).

Next, the system uses a criterion $w(\mathbf{X}) > (w_I + w_R)/2$ to decide whether to (try to) solve the constraints in linearized form. If not, the system goes to step 744 (step 731). Note that, w_I denotes the width of the smallest box for which

25 $\mathbf{M}^I = \mathbf{B}\mathbf{J}(\mathbf{x}, \mathbf{X})$ is irregular. If \mathbf{M}^I is regular for a given box, w_R denotes the width of the largest box for which \mathbf{M}^I has been found to be regular (step 731).

The system then replaces $n-r$ of the variables by their interval bounds, and renames the remaining variables as x_1, \dots, x_r . Then the system linearizes the constraint functions as functions of the variables now called x_1, \dots, x_r . The system also computes an approximate inverse \mathbf{B} of the approximate center of the Jacobian

5 **J(x,X)** (step 732).

Next, the system preconditions the linearized system, and if the preconditioned coefficient matrix is regular, the system finds the hull of its solution. In the process, regularity of the preconditioned coefficient matrix is determined. If the result is empty, the system goes to step 703 (step 733).

10 If $w[f(\mathbf{X})] \leq \varepsilon_F$ and $w(\mathbf{X}) \leq \varepsilon_X$, the system puts \mathbf{X} in list L_2 and goes to step 703 (step 734).

Next, the system analytically multiplies the nonlinear system of constraint equations by the matrix \mathbf{B} . It does so without replacing any variables by their interval bounds so that appropriate combinations and cancellations can be made.

15 After the analytic multiplication is complete, the system replaces the fixed variables (as chosen in step 732) by their interval bounds. The system also applies term consistency to solve the i -th nonlinear equation of the preconditioned nonlinear system for the i -th variable (as renamed in step 732) for $i=1, \dots, r$. If the result is empty, the system goes to step 703 (step 735).

20 If $w[f(\mathbf{X})] \leq \varepsilon_F$ and $w(\mathbf{X}) \leq \varepsilon_X$, the system puts \mathbf{X} in list L_2 and goes to step 703 (step 736).

Next, the system applies term consistency to solve the i -th nonlinear equation of the preconditioned nonlinear system for the i -th (renamed) variable for $i=1, \dots, r$. If the result is empty, the system goes to step 703. If the existence of a feasible point is proved (see Section 10.12), the system uses the result to update f_bar (step 737).

If $w[f(\mathbf{X})] \leq \varepsilon_F$ and $w(\mathbf{X}) \leq \varepsilon_X$, the system puts \mathbf{X} in list L_2 and goes to step 703 (step 738).

Next, the system applies box consistency to solve the i -th nonlinear equation of the preconditioned nonlinear system for the i -th (renamed) variable for 5 $i=1,\dots,r$. If the result is empty, the system goes to step 703 (step 739).

If $w[f(\mathbf{X})] \leq \varepsilon_F$ and $w(\mathbf{X}) \leq \varepsilon_X$, the system puts \mathbf{X} in list L_2 and goes to step 703 (step 740).

Next, the system uses the criterion $w(\mathbf{X}) > (w_I + w_R)/2$ to decide whether to (try to) solve the John conditions using a Newton method. If not, the system goes 10 to step 744 (step 741).

The system then applies one step of the interval Newton method for solving the John conditions. If the result is empty, the system goes to step 703. If the existence of a solution of the John conditions is proved, the system updates f_bar (step 742).

15 If the box $\mathbf{X}^{(1)}$ (as defined in step 703) has been sufficiently reduced, the system puts \mathbf{X} in L_1 and goes to step 703 (step 743).

Any previous step that used term consistency, a Newton step, or a Gauss-Seidel step might have generated gaps in the interval components of \mathbf{X} . The system merges any such gaps when possible, and splits the box. This may involve 20 deleting some gaps. The system places the subboxes (generated by splitting) in the list L_1 and goes to step 703 (step 744).

If the list L_2 is empty, the system indicates that there is no feasible point in $\mathbf{X}^{(0)}$. The system then goes to step 750 (step 745).

25 If $f_bar < +\infty$ and there is only one box in L_2 , the system goes to step 750 (step 746).

For each box \mathbf{X} in L_2 , if $\sup(f(m(\mathbf{X}))) < f_bar$, the system tries to prove existence of a feasible point. The system uses the results to update f_bar (step 747).

Next, the system deletes any box \mathbf{X} from L_2 for which the lower bound of
 5 $f(\mathbf{X}) > f_bar$ (step 748).

The system denotes the remaining boxes by $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(s)}$ where s is the number of boxes remaining in L_2 . The system then determines

$$\underline{F} = \min_{1 \leq i \leq s} \underline{f}(\mathbf{X}^{(i)}) \quad \text{and} \quad \overline{F} = \max_{1 \leq i \leq s} \overline{f}(\mathbf{X}^{(i)}) \text{ (step 749).}$$

10

Finally, the system terminates (step 750).

If, at termination, the list L_2 is empty, then all of the initial box $\mathbf{X}^{(0)}$ has been eliminated. This provides proof that the initial box $\mathbf{X}^{(0)}$ does not contain a feasible point.

15

Assume that at least one box remains in the list L_2 . What we have proved in this case depends on the final value of f_bar . If $f_bar < +\infty$, then we know that a feasible point exists in the initial box $\mathbf{X}^{(0)}$. If $f_bar = \infty$, there may or may not be a feasible point in $\mathbf{X}^{(0)}$.

20

Consider the case $f_bar < \infty$. No matter how poor the bound f_bar on f^* , we know that a global solution exists in $\mathbf{X}^{(0)}$; and it is in one of the remaining boxes. Also, we know that,

$$\underline{F} \leq f^* \leq \overline{F}.$$

25

If only one box \mathbf{X} remains, then it must contain the global solution. In this case,

$$\underline{f}(\mathbf{X}) \leq f^* \leq \min\{\overline{f}(\mathbf{X}), f_bar\} \text{ and } \overline{f}(\mathbf{X}) - \underline{f}(\mathbf{X}) \leq \varepsilon_F.$$

Therefore, $f(\mathbf{x}) - f^* \leq \varepsilon_F$ for every point \mathbf{x} in the box. Also,

5 $x^* - \underline{X}_i \leq \varepsilon_X \text{ and } \overline{X}_i - x^* \leq \varepsilon_X \quad (i = 1, \dots, n).$

If more than one box remains, it is possible that one contains a local solution at which f is less than our upper bound f_bar . Also, there might be more than one global solution occurring in separate boxes. We know only that

10 $\underline{F} \leq f^* \leq \min\{\overline{F}, f_bar\}$

and that the global minimum point(s) are in the remaining boxes.

If the final value of f_bar is ∞ and $\mathbf{X}^{(0)}$ is not entirely deleted, then $\mathbf{X}^{(0)}$
 15 may or may not contain a feasible point. We do not know. It is highly probable that a solution exists since, otherwise, we expect all of $\mathbf{X}^{(0)}$ to be deleted. However, we do know that if a feasible point does exist in $\mathbf{X}^{(0)}$, then

20 $\underline{F} \leq f^* \leq \overline{F}$

and x^* is somewhere in the remaining box(es). All local solutions in the initial box are contained in the final solution box(es).

It is possible that every point in the initial box $\mathbf{X}^{(0)}$ is infeasible. However, our process can delete all of $\mathbf{X}^{(0)}$ (and thus prove there is no solution) only if every
 25 point in $\mathbf{X}^{(0)}$ is *proved* to be feasible (i.e., is certainly infeasible). Even if every point in $\mathbf{X}^{(0)}$ is certainly infeasible, our process may still not delete all of $\mathbf{X}^{(0)}$. The probability of it doing so in this case is greater when the box size tolerance ε_X is smaller.

Thus, we can prove there is no feasible point in $\mathbf{X}^{(0)}$, but we do not guarantee doing so when this is, in fact, the case.

Regardless of what procedure is used to delete a part of $\mathbf{X}^{(0)}$, we know that the deleted part cannot contain a solution.

5 The foregoing descriptions of embodiments of the present invention have been presented only for purposes of illustration and description. They are not intended to be exhaustive or to limit the present invention to the forms disclosed.

For example, global optimization problems can contain both inequality and equality constraints. In this case, the present invention can be combined with
10 a system that solves an interval inequality constrained global optimization problem to form an overall system.

In another example, although the present invention describes the use of derivatives in certain situations, it is often possible to use slopes instead of derivatives.

15 Accordingly, many modifications and variations will be apparent to practitioners skilled in the art. Additionally, the above disclosure is not intended to limit the present invention. The scope of the present invention is defined by the appended claims.